# An Interpolation Curve using a Spline in Tension

by

Daniel G. Schweikert



ABSTRACT

The use of a linearized mathematical spline for interpolation between given points occasionally yields extraneous inflection points which for some applications, notably the fairing of ship lines, are not acceptable. To avoid this difficulty, a spline in tension is considered and the value of tension which is sufficient to remove all the extraneous inflection points is determined. The slopes of the interpolating curve at each of the given stations are considered as the unknowns and can be calculated without the necessity of numerical matrix inversion.

## I. INTRODUCTION

Frequently in engineering work one encounters the need to relate a smooth curve to a large number of discrete points. Where the discrete points represent experimental data, one usually desires a smooth curve which will approximate the given data; thus, the curve is not required to pass through every given point. However, when the discrete points are intended to indicate the geometric outline of a physical body, the smooth curve is usually required to pass through the exact locations of all the given points; such a curve serves the purpose of interpolating points between the given points. Only the second type of curve, the interpolation curve, will be considered here.

Historically, draftsmen have determined a "smooth", or "faired", interpolation curve with the aid of a long flexible beam or "spline". The spline is constrained to pass through the given points in a plane, and heavy objects - "dogs" - placed between the given points keep the spline from moving while its deflection curve is being outlined on the drawing. Although seemingly not subject to variations, two draftsmen will seldom achieve identical results. In spite of these slight variations the method, in general, yields highly satisfactory results.

The usual mathematical analysis of the draftman's spline considers the small deflections of a simply supported elastic beam. Using "Strength of Materials" assumptions, the deflection curve can be represented by cubic polynomials

connecting the given points with adjacent cubics having their first and second derivatives equal at their common given point. The method has been aptly termed a "piecewise polynomial" method by Birkhoff and Garabedian [1]; a derivation of the form used by them for one variable (the slopes at the given points are taken as the unknowns) appears in Part II. Although Schoenberg notes [2] related work on piecewise polynomial approximation found in actuarial literature on "osculatory interpolation", he is apparently the first to extensively treat the polynomial analog of the draftsman's spline [3,4].

The cubic polynomial "spline" curve has many useful properties [5,6,7] and has been successfully used as an interpolating curve in many problems far removed from those treated by its physical counterpart, the draftsman's spline. Unfortunately, the cubic polynomial method is not satisfactory in the problem closest to its origin: producing a smooth geometrical outline which passes through the given points. The difficulty centers around certain conditions on curvature implied by the physical problem. One may require a convex interpolation curve for a "convex" set of given points. For instance, consider the five points in Figure 1; the interior points 2, 3, and 4 have positive second differences (see eq. 22) and it would be reasonable to require the interpolating curve to have a positive second derivative. Such a requirement is indispensable where the interpolation curve represents the geometrical outline of a form which must be fabricated, such as a ship's hull or automobile

body. The cubic polynomial method does not always satisfy this plausible requirement; under conditions discussed in Part 2.2, the curve will have a negative second derivative in certain regions where a positive second derivative is expected from the location of the given points, thus introducing unwanted, or extraneous, inflection points.

Customarily, the draftsman's results do not have the extraneous inflection points predicted by cubic polynomial representation. Thus, by refining the assumptions made to obtain the cubic polynomial representation, we may possibly obtain a better mathematical representation which duplicates the physical model's lack of extraneous inflection points. The assumptions can be considered as follows: 1) the group of "Strength of Materials" assumptions concerning the distribution of stress and strain in the beam; 2) small deflections, or more precisely, curvature taken equal to the second derivative; 3) simple supports at the given points. The validity of the first group of assumptions seems certain, at least with respect to extraneous inflection points, since experiment or physical intuition can quickly supply examples of simply supported physical splines which have extraneous inflection points under approximately the same conditions as those which yield such points in the cubic polynomial method. The second assumption can be dismissed from consideration since it will be shown that the conditions under which an extraneous inflection point occurs in the cubic polynomial method is independent of the magnitude of the

deflections. The assumption of simple supports, however, is immediately open to question. The heavy "dogs" provide the physical mechanism which could produce reaction moments and maintain tension in the spline. It seems likely that the draftsman recognizes the "convexity" of the given points and manipulates the spline by deviating from simple support conditions in order to achieve the implied curvature conditions, thereby avoiding extraneous inflection points.

The question now arises as to what parts of the simple support assumption can be successfully retained. Since only a small amount of tension applied to a simply supported spline would physically remove the extraneous inflection points in most cases, one is led to consider the simpler condition of tension alone and neglect the reaction moments.

The value of tension can be expected to change abruptly at each of the "dogs". However, as a mathematical expediency, tension will be considered constant over the length of the spline. This assumption permits the use of a mathematical model in which the tension is produced by axial loads at the end of the spline rather than by reaction forces at the support points. Thus, the net result of these changes will be the analysis, in Part IV, of a simply supported spline in tension from a "Strength of Materials" viewpoint.

# II. SIMPLY SUPPORTED SPLINE

## 2.1 Basic Equations

Let the ordinates $y_o, y_1, \ldots, y_{n+1}$ be given at $x_o, x_1, \ldots, x_{n+1}$ respectively. For convenience consider the unit spacing $x_{i+1} - x_i = 1$ $(i = 0, \ldots, n)$. Let $y(x)$ be an interpolation curve through these points and define $y_i'$ and $y_i''$ as the first and second derivatives of $y(x)$ at $x_i$ for $i = 0, 1, \ldots, n+1$. The interpolation curve associated with the deflection curve of a simply supported elastic beam (spline) is made up of a sequence of cubic polynomials $f_i(x)$ such that

$$\begin{aligned} y(x) &= f_i(x) \quad \text{for} \quad x_i \leq x < x_{i+1} \quad (i = 0, \ldots, n) \\ y(x_{n+1}) &= \lim_{x \to x_{n+1}} f_n(x) \end{aligned} \tag{1}$$

with the following conditions on value, slope, and second derivative:

$$f_i(x_i) = y_i \qquad i = 0, \ldots, n, \tag{2}$$

$$\lim_{x \to x_i} f_{i-1}(x) = y_i \qquad i = 1, \ldots, n+1, \tag{3}$$

$$\lim_{x \to x_i} f_{i-1}'(x) = f_i'(x_i) \qquad i = 1, \ldots, n, \tag{4}$$

$$\lim_{x \to x_i} f_{i-1}''(x) = f_i''(x_i) \qquad i = 1, \ldots, n. \tag{5}$$

The cubic, $f_i(x)$, can be expressed in terms of the ordinates and slopes at each of its end points, $x_i$ and $x_{i+1}$:

$$f_i(x) = (x_{i+1}-x)y_i + (x-x_i)y_{i+1} +$$
$$(x_{i+1}-x)(x-x_i)\{[y_i' - (y_{i+1}-y_i)](x_{i+1}-x) \qquad (6)$$
$$- [y_{i+1}' - (y_{i+1}-y_i)](x-x_i)\}$$
$$\text{for} \quad x_i \leq x < x_{i+1} \quad (i = 0,\ldots,n).$$

The form of $f_i(x)$ shown in (6) identically satisfies the conditions on value and slope expressed in (2), (3) and (4); however, it does not, in general, satisfy the continuity condition on the second derivative expressed in (5). In order to expand (5) we first differentiate (6) twice:

$$f_i''(x) = y_{i+1}' - y_i' + 3[2(y_{i+1}-y_i) - (y_{i+1}'+y_i')][x_{i+1}-2x+x_i]$$
$$\text{for} \quad x_i \leq x < x_{i+1} \quad (i = 0,\ldots,n). \qquad (7)$$

From (7),

$$f_i''(x_i) = y_{i+1}'-y_i'+3[2(y_{i+1}-y_i) - (y_{i+1}'+y_i')] \quad (i = 0,\ldots,n) \quad (8)$$

and

$$\lim_{x \to x_i} f_{i-1}''(x) = y_i'-y_{i-1}'-3[2(y_i-y_{i-1})-(y_i'+y_{i-1}')] \quad (i=1,\ldots,n+1). \quad (9)$$

Using (8) and (9), (5) becomes

$$y_{i-1}' + 4y_i' + y_{i+1}' = 3(y_{i+1} - y_{i-1}) \quad (i = 1,\ldots,n). \quad (10)$$

Equation (10) gives n equations for the n+2 unknowns $y_i'$ ($i = 0, \ldots, n+1$) necessary to determine $y(x)$ throughout $x_0 \leq x \leq x_{n+1}$ (cf. (1) and (6)). The two additional required equations are usually obtained by specifying end conditions on slope or second derivative at $x_0$ and $x_{n+1}$.

If the end slopes are specified, the number of unknowns is immediately reduced to n and these will be determined by (10) alone.

If the second derivative $y_0''$ is specified instead of $y_0'$ then the equation

$$y_0'' = f_0''(x_0), \tag{11}$$

or, using (7) to expand (11) into a more convenient form,

$$2y_0' + y_1' = -\tfrac{1}{2}y_0'' + 3(y_1-y_0) \tag{12}$$

is added to eqs. (10) to provide n+1 equations for the n+1 unknowns, $y_i'$ ($i = 0,\ldots,n$). Likewise, the equation

$$y_{n+1}'' = \lim_{x \to x_{n+1}} f_n''(x), \tag{13}$$

or, using (7) to expand (13),

$$y_n' + 2y_{n+1}' = \tfrac{1}{2}y_{n+1}'' + 3(y_{n+1} - y_n), \tag{14}$$

is considered with eqs. (10) when $y_{n+1}''$ is specified instead of $y_{n+1}'$. If the physical spline is simply supported at the end-points, the internal moment at those points is zero. In this idealized representation of the physical spline, the specification of $y_0'' = y_{n+1}'' = 0$ is equivalent to having zero end moments and is the most frequently used end condition.

## 2.2 Conditions for an Inflection Point

Assume that an interpolation curve, $y(x)$, based on the above cubic polynomial representation of a spline, has been determined. We wish to determine the conditions which are

associated with the appearance of inflection points. Since $f_i(x)$ is specified by its endpoint values $y_i$, $y_{i+1}$ and slopes $y_i'$, $y_{i+1}'$, the analysis of one interval will yield results typical of all intervals. Consider the interval $x_i \leq x \leq x_{i+1}$. Since $f_i''(x)$ (cf. (7)) is linear there can be an inflection point if and only if $y_i''$ and $y_{i+1}''$ are of opposite sign. In view of the discussion in Part I, we must study the second differences associated with $x_i$ and $x_{i+1}$ in order to determine, assuming $y_i''$ and $y_{i+1}''$ are of opposite sign, if such an inflection point is extraneous. Before doing this it will be helpful to discuss the inflection point conditions in terms of slopes rather than second derivatives.

Let $s(x)$ be the difference between the slope of $y(x)$ and a polygonal function connecting the given points. Then, in the interval being considered,

$$s(x) = y'(x) - (y_{i+1} - y_i) \qquad x_i \leq x \leq x_{i+1}, \tag{15}$$

with its values at $x_i$ and $x_{i+1}$ being

$$s_i = y_i' - (y_{i+1} - y_i) \tag{16}$$

and

$$s_{i+1} = y_{i+1}' - (y_{i+1} - y_i), \tag{17}$$

respectively. Substituting (16), (17) in (8) and (9), the second derivatives $y_i''$ and $y_{i+1}''$ can be expressed in terms of $s_i$ and $s_{i+1}$:

$$y_i'' = f_i''(x_i) = s_{i+1} - s_i - 3(s_{i+1} + s_i) \tag{18}$$

$$y_{i+1}'' = \lim_{x \to x_{i+1}} f_i(x) = s_{i+1} - s_i + 3(s_{i+1} + s_i) \tag{19}$$

Defining

$$\zeta_i \equiv \frac{s_i}{s_{i+1}} , \tag{20}$$

the ratio of second derivates has the form

$$\frac{y_i''}{y_{i+1}''} = -\frac{1+2\zeta_i}{2+\zeta_i} ; \tag{21}$$

this ratio is positive if and only if $-2 < \zeta_i < -\frac{1}{2}$. If $s_i$ and $s_{i+1}$ have the same sign ($\zeta_i > 0$) we would expect an inflection point. For $s_i$ and $s_{i+1}$ of opposite sign an inflection point occurs if $s_i$ and $s_{i+1}$ differ in magnitude by more than a factor of two; an inflection point in this case is not usually desirable. For example, consider the single interval problem ($n = 0$) shown in Figure 2. With $y_o = 0$, $y_1 = \frac{1}{4}$, and specified and slopes $y_o' = 1$, $y_1' = 0$, we have, from (16), (17) and (20), $s_o = \frac{3}{4}$, $s_1 = -\frac{1}{4}$ and $\zeta_o = -3$. The cubic polynomial interpolation curve, represented by the solid line, shows an extraneous inflection point and subsequent reverse in curvature near the right end. Clearly, a curve such as the broken curve in Figure 2 has the desired smoothness which the solid curve lacks, even though its second derivative must have a larger maximum and could be considered less "smooth" in a mathematical sense. This situation suggests that some additional condition on curvature would be useful to further define, mathematically, the desired type of smoothness.

## III. CONDITIONS ON CURVATURE.

In order to avoid extraneous inflection points, the discussion concerning Figure 1 leads to the requirement that $y_i''$ have the same sign as the second difference at $x_i$, which will be denoted as $d_i$. Although this brief statement is intuitively plausible, it leaves two definitions and some intervening arguments in need of elaboration.

The first problem is that the usual definition of the second difference at $x_i$,

$$d_i = (y_{i+1}-y_i) - (y_i-y_{i-1}) = y_{i+1} - 2y_i + y_{i-1} \quad (i = 1,\ldots,n) \tag{22a}$$

leaves us without an indicator of the desired sign of $y_i''$ at the endpoints $x_o$ and $x_{n+1}$. If $y_o''$ is specified then the sign is obviously the desired one, however, for notational convenience in stating the curvature conditions, we formally define:

$$d_o = y_o'', \quad \text{if } y_o'' \text{ is specified,}$$

and similarly (22b)

$$d_{n+1} = y_{n+1}'', \quad \text{if } y_{n+1}'' \text{ is specified.}$$

For the case of a specified end slope we look again to Figure 1. If the specified slope at 1 is that of the dashed line a, we would expect a positive second derivative at 1; if the slope is that of b then a negative second derivative is desirable. A convenient indicator which duplicates these results is a second difference which uses the specified slope as the leading first difference; i.e.

$$d_o = (y_1 - y_o) - y_o', \quad \text{if } y_o' \text{ is specified,}$$

and similarly (22c)

$$d_{n+1} = y_{n+1}' - (y_{n+1} - y_n), \quad \text{if } y_{n+1}' \text{ is specified.}$$

Definition. Assume that $y(x)$ is an interpolation curve of class $C^2[x_o, x_{n+1}]$ through the given points $y_i$ at $x_i$ ($i = 0, \ldots, n+1$), satisfying conditions on slope or second derivative at $x_o$ and $x_{n+1}$. Assume further that $y(x)$ has at most one inflection point in the interval $x_i \leq x \leq x_{i+1}$ ($i = 0, \ldots, n$). Then, such an inflection point, if it exists, will be extraneous if $d_i$ and $d_{i+1}$ have the same sign.

Theorem 1. Assume $y(x)$ as in the above definition. Then the condition that $y_i''$ and $d_i$ ($i = 0, \ldots, n+1$) have the same sign is necessary and sufficient for $y(x)$ to have no extraneous inflection points.

Proof. The condition is clearly sufficient since for every interval such that $d_i$ and $d_{i+1}$ have the same sign, $y_i''$ and $y_{i+1}''$ would have the same sign and there would be no inflection point. Proof of necessity is more involved and will be demonstrated with the aid of Figure 3 and the notation that, say, $(+,-)_i$ will mean $y_i''$ is positive and $d_i$ is negative. Assume that there is some point $i$ such that $y_i''$ and $d_i$ have the same sign, say $(+,+)_i$, and $y_{i+1}''$ and $d_{i+1}$ have opposite signs. We have two choices, $(-,+)_{i+1}$ or $(+,-)_{i+1}$; the former implies an extraneous inflection point so we choose the latter. Because of the signs of $d_i$ and $d_{i+1}$, it is clear from Figure 3 that $y_{i+2}''$ must be

negative since only one inflection point in an interval is permitted; then $(-,+)_{i+2}$ is the only combination that doesn't make the inflection point extraneous. Repeating the argument for i+3, i+4,... we find that the alternation of signs must continue if extraneous inflection points are to be avoided. In this case, however, y(x) cannot satisfy the final boundary condition. For instance, if i+4 is the last point, the dashed line would be a typical specified slope associated with the positive $d_{i+4}$, a value impossible for $y'_{i+4}$ to achieve. If $y''_{i+4}$ had been specified the contradiction is more direct since $d_{i+4}$ has the same sign by definition. The initial assumption that there is at least one point i such that $d_i$ and $y''_i$ have the same sign, if there are no extraneous inflection points, can be shown by an argument similar to the one just completed. If one assumes $d_i$ and $y''_i$ have opposite signs at all points then the boundary condition at either the first or last point cannot be satisfied.

The cubic interpolation curve may satisfy the above conditions on the second derivative (i.e., $y''_i$ and $d_i$ have the same sign), but if it does not, nothing can be changed to correct it. In order to find a spline curve which has a parameter that can be varied to satisfy the curvature conditions, we are led, with some physical justification as well, to consider a simply supported spline in tension.

## IV. SIMPLY SUPPORTED SPLINE IN TENSION

### 4.1 Basic Equations

For the spline in tension we have, corresponding to (6), the form [8]:

$$f_i(x) = (x_{i+1}-x)y_i + (x-x_i)y_{i+1} +$$

$$+ \frac{1}{\eta^2-1}[y_i' + \eta y_{i+1}' - (\eta+1)(y_{i+1}-y_i)][\frac{\sinh p(x-x_i)-(x-x_i)\sinh p}{\sinh p - p}]$$

$$- \frac{1}{\eta^2-1}[\eta y_i'+y_{i+1}'-(\eta+1)(y_{i+1}-y_i)][\frac{\sinh p(x_{i+1}-x)-(x_{i+1}-x)\sinh p}{\sinh p - p}], \tag{23}$$

$$\text{for} \quad x_i \le x < x_{i+1} \quad (i = 0,\ldots,n),$$

where

$$\eta = \frac{p\cosh p - \sinh p}{\sinh p - p} \tag{24}$$

and

$$P \propto \sqrt{\text{tension}}$$

Although not immediately obvious, (23) reduces to (6) for $p = 0$. Figure 4 shows a graph of $\eta$ as a function of p. For p large, $\eta$ approaches p-1 asymptotically.

Corresponding to (7), the second derivative is

$$f_i''(x) = \frac{1}{\eta^2-1}\left\{[y_i' + \eta y_{i+1}' - (\eta+1)(y_{i+1}-y_i)]\frac{p^2\sinh p(x-x_i)}{\sinh p - p}\right.$$

$$\left. - [\eta y_i' + y_{i+1}' - (\eta+1)(y_{i+1}-y_i)]\frac{p^2\sinh p(x_{i+1}-x)}{\sinh p - p}\right\} \tag{25}$$

$$\text{for} \quad x_i \le x < x_{i+1} \quad (i = 0,\ldots,n).$$

From (25),

$$f_i''(x_i) = \frac{-1}{\eta^2-1}[\eta y_i'+y_{i+1}'-(\eta+1)(y_{i+1}-y_i)] \frac{p^2 \sinh p}{\sinh p - p} \tag{26}$$

and

$$\lim_{x \to x_{i+1}} f_{i-1}''(x) = \frac{1}{\eta^2-1}[y_{i-1}'+\eta y_i'-(\eta+1)(y_i-y_{i-1})] \frac{p^2 \sinh p}{\sinh p - p} \tag{27}$$

As in the cubic polynomial representation of the spline, the form of $f_i(x)$, (23), assures that the conditions on value and slope, expressed in (2), (3) and (4), are identically satisfied. Using (26) and (27), the continuity condition on the second derivative, (5), becomes

$$y_{i-1}' + 2\eta y_i' + y_{i+1}' = (\eta+1)(y_{i+1}-y_{i-1}) \quad (i = 1,\ldots,n) \tag{28}$$

Note that since $\eta$ equals 2 in the limit as p approaches zero, (28) reduces to (10) as $p \to 0$.

Equation (28) gives n equations for the n+2 unknowns $y_i$, $(i = 0,1,\ldots,n+1)$. If the end slopes $y_0'$ and $y_{n+1}'$ are specified, there are only n unknowns and equations (28) are sufficient. If the second derivatives at the ends, $y_0''$ and $y_{n+1}''$ are specified, then the equations (cf. (12), (14))

$$\begin{aligned} \eta y_0' + y_1' &= -y_0'' \frac{\sinh p - p}{p^2 \sinh p}(\eta^2-1) + (\eta+1)(y_1-y_0) \\ y_n'+\eta y_{n+1}' &= y_{n+1}'' \frac{\sinh p - p}{p^2 \sinh p}(\eta^2-1) + (\eta+1)(y_{n+1}-y_n), \end{aligned} \tag{29}$$

are combined with (28) to provide n+2 equations for the n+2 unknowns.

4.2. Conditions for an Inflection Point.

Combining (25) with (26) and (27), the second derivative can be expressed in terms of its values at the interval end-points:

$$f_i''(x) = y_i'' \frac{\sinh p(x_{i+1}-x)}{\sinh p} + y_{i+1}'' \frac{\sinh p(x-x_i)}{\sinh p} \tag{30}$$

$$x_i \leq x \leq x_{i+1} \qquad (i = 0,\ldots,n)$$

Inspection of this form shows that although $f_i''(x)$ is not a linear function as in the cubic polynomial case, it still has the property that it will not cross zero unless the end values, $y_i''$ and $y_{i+1}''$, have opposite signs. For instance, if $y_i''$ and $y_{i+1}''$ are both positive then, going from $x_i$ to $x_{i+1}$, the first term in (30) decreases monatonically from $y_i''$ to zero while the second term increases monatonically from zero to $y_{i+1}''$; thus both terms are positive in the interval and there can be no inflection point. A similar argument shows that if $y_i''$ and $y_{i+1}''$ are of opposite sign then there is one and only one inflection point. Thus, we have shown that the interpolation curve, $y(x)$ (cf. (1) - (5)), using the hyperbolic spline representation $f_i(x)$ shown in (23), satisfies the assumptions on $y(x)$ specified in the Definition and Theorem 1 discussed in Part III.

The inflection point conditions in terms of $s_i$ are developed in same manner as that discussed in Part 2.2 for the cubic polynomial representation (see (15) - (21)). The final results are that there will <u>not</u> be an inflection point in the interval $x_i < x < x_{i+1}$ if and only if $-\eta < \zeta_i < -1/\eta$

and that for $s_i$ and $s_{i+1}$ of opposite sign, an extraneous inflection point will not occur if $s_i$ and $s_{i+1}$ differ in magnitude by less than a factor of $\eta$. At this point one may be led to the conclusion that by increasing tension sufficiently, $\eta$ will accommodate any specified ratio $\zeta_i$ and therefore all extraneous inflection points can be removed. The problem is that $s_i$ is a function of $\eta$ and thus as $\eta$ changes, so does $\zeta_i$. Since the hyperbolic spline representation approaches a polygonal function as $\eta$ becomes large, the values of $\zeta_i$ approach finite limits, which places the conclusion on firmer ground. It is more useful, however, to establish this conclusion by completely determining the effect of tension on slope and curvature throughout all the intervals. Clearly, an explicit solution for the unknowns $y'_i$ $(i = 0,\ldots,n+1)$ is desirable.

## 4.3 An Explicit Solution.

Consider first the case of given $y'_o$ and $y'_{n+1}$. Equations (28) can be written in matrix form,

$$N\,Y' = C \tag{31}$$

with

$$N = \begin{bmatrix} 2\eta & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 2\eta & 1 & & & & 0 \\ 0 & 1 & 2\eta & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & & & \ddots & 2\eta & 1 & 0 \\ 0 & & & & 1 & 2\eta & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 1 & 2\eta \end{bmatrix} \tag{32}$$

$$Y' = \begin{bmatrix} y_1' \\ y_2' \\ \cdot \\ \cdot \\ \cdot \\ y_{n-1}' \\ y_n' \end{bmatrix} \quad (33) \qquad C = \begin{bmatrix} (\eta+1)(y_2-y_o) - y_o' \\ (\eta+1)(y_3-y_1) \\ (\eta+1)(y_4-y_2) \\ \cdot \\ \cdot \\ \cdot \\ (\eta+1)(y_{n-1}-y_{n-3}) \\ (\eta+1)(y_n-y_{n-2}) \\ (\eta+1)(y_{n+1}-y_{n-1}) - y_{n+1}' \end{bmatrix} \quad (34)$$

Since $\eta \geq 2$ for all $p \geq 0$, the diagonal element in each row of N is larger than the sum of the absolute values of the remaining elements in that row; thus, by a well known theorem on determinants [9], N has an inverse. In fact, the inverse of N (say, order m) can be written explicitly [10,11]:

$$N^{-1} = \frac{1}{P_m} \begin{bmatrix} +P_o P_{m-1} & -P_o P_{m-2} & +P_o P_{m-3} & \cdots & (-1)^{m+1} P_o P_o \\ -P_o P_{m-2} & +P_1 P_{m-2} & -P_1 P_{m-3} & \cdots & (-1)^m P_1 P_o \\ +P_o P_{m-3} & -P_1 P_{m-3} & +P_2 P_{m-3} & \cdots & (-1)^{m+1} P_2 P_o \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^{m+1} P_o P_o & (-1)^m P_1 P_o & (-1)^{m+1} P_2 P_o & \cdots & P_{m-1} P_o \end{bmatrix} \quad (35)$$

where

$$\begin{aligned} P_o &= 1 \\ P_1 &= 2\eta \\ P_2 &= 2\eta P_1 - P_o \\ P_3 &= 2\eta P_2 - P_1 \\ P_4 &= 2\eta P_3 - P_2 \\ &\vdots \\ P_m &= 2\eta P_{m-1} - P_{m-2} \end{aligned} \quad (36)$$

or, since $P_k$ is the Tschebyscheff polynomial of the second kind, $U_k(\eta)$, the direct representation is

$$P_i = 2^i \prod_{k=1}^{i} (\eta - \cos\frac{k\pi}{i+1}) \quad (i = 1,\ldots,m)$$
$$P_o = 1. \qquad (37)$$

Defining a matrix R as the adjoint of N,

$$R = P_m N^{-1}, \qquad (38)$$

equation (35) can be written in indicial form:

$$R_{ij} = (-1)^{i+j} P_{i-1}P_{m-j} \quad i \leq j \quad i,j = 1,\ldots,m$$
$$R_{ji} = R_{ij} \qquad i,j = 1,\ldots,m \qquad (38')$$

Also we note symmetry about the non-principal diagonal:

$$R_{m+1-j\ m+1-i} = R_{ij} \quad i,j = 1,\ldots,m$$

In the above case of given end slopes, m = n. In the case of given $y_o''$ and $y_{n+1}''$ given, N takes on a slightly changed form since equations (29) must be included:

$$N = \begin{bmatrix} \eta & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 2\eta & 1 & & & & 0 \\ 0 & 1 & 2\eta & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & & & \ddots & 2\eta & 1 & 0 \\ 0 & & & & 1 & 2\eta & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 1 & \eta \end{bmatrix} \qquad (39)$$

with order m = n+2. The inverse of this second form of N, is identical to the form shown in (35) if $P_i$ is redefined as [12,11]:

$$\begin{aligned} P_o &= 1 \\ P_1 &= \eta \\ P_2 &= 2\eta P_1 - P_o \\ P_3 &= 2\eta P_2 - P_1 \\ &\vdots \\ P_{m-1} &= 2\eta P_{m-2} - P_{m-3} \\ P_m &= \eta P_{m-1} - P_{m-2} \end{aligned} \tag{40}$$

Except for $P_m$, the $P_i$ $(i = 0,\ldots,m-1)$ in (40) are Tschebyscheff polynomials of the first kind, $T_i(\eta)$, and have adirect representation

$$P_o = 1$$

$$P_i = 2^{i-1} \prod_{k=1}^{i} \left(\eta - \cos \frac{2k-1}{2i} \pi\right) \qquad (i = 1,\ldots,m-1)$$

With $y''_o$ and $y''_{n+1}$ specified, the column vectors Y' and C (cf. (31)) that correspond to N (cf. (39)) are:

$$Y'' = \begin{bmatrix} y'_o \\ y'_1 \\ \cdot \\ \cdot \\ \cdot \\ y'_{n+1} \end{bmatrix} \quad (42) \qquad C = \begin{bmatrix} (\eta+1)(y_1-y_o) - (\eta^2-1)\dfrac{\sinh p - p}{p^2 \sinh p}\, y''_o \\ (\eta+1)(y_2-y_o) \\ (\eta+1)(y_3-y_1) \\ \vdots \\ (\eta+1)(y_{n+1}-y_{n-1}) \\ (\eta+1)(y_{n+1}-y_n) + (\eta^2-1)\dfrac{\sinh p - p}{p^2 \sinh}\, y''_{n+1} \end{bmatrix}$$

(43)

Because of the similarity in the forms of the solution, the case of given $y''_0$ and $y''_{n+1}$ will not be discussed further.

Theorem 2. Let an interpolation curve y(x) be of the form specified in eqs. (1) - (5) with $f_i(x)$ $(i = 0,\ldots,n)$ being the hyperbolic spline representation given in (23). Assume that the end slopes $y'_0$ and $y'_{n+1}$ are specified and that $d_i \neq 0$ $(i = 1, \ldots,n)$. (For the case of $d_i = 0$, i.e., three given points in a line, see Part 4.4.) Then, the condition expressed in (5) and equivalently in (28) and in (31) has the solution

$$Y' = N^{-1}C \tag{44}$$

where $N^{-1}$ is given by (35). Thus, y(x) is completely determined.

Further, there exists a value of $\eta \geq 2$ such that the curvature conditions expressed in Theorem 1 are satisfied, i.e.

$$y''_i\, d_i > 0 \qquad (i = 0,\ldots,n+1). \tag{45}$$

Proof. First we will show that (35) is the correct inverse. Using (38), this is equivalent to proving.

$$NR = P_m I \tag{46}$$

Let

$$B = NR \tag{47}$$

then using (32) and (38') the diagonal terms can be expressed: $(i \neq 1,m)$

$$\begin{aligned} B_{ii} &= (-P_{i-2}P_{m-i} + 2\eta P_{i-1}P_{m-1} - P_{i-1}P_{m-i-1}) \\ &= P_{i-1}(2\eta p_{m-i} - P_{m-i-1}) - P_{i-1}P_{m-i}. \end{aligned} \tag{48a}$$

Using (36) repeatedly:

$$B_{ii} = P_{i-1}\, P_{m-i+1} - P_{i-2}\, P_{m-i}$$

$$= (2\eta P_{i-2} - P_{i-3})P_{m-i+1} - P_{i-2}P_{m-i}$$

$$= P_{i-2}(2\eta P_{m-i+1} - P_{m-i}) - P_{i-3}P_{m-i+1}$$

$$= P_{i-2}P_{m-i+2} - P_{i-3}\, P_{m-i+1}$$

$$\vdots$$

$$= P_1 P_{m-1} - P_o\, P_{m-2}$$

$$= P_o\, 2\eta P_{m-1} - P_o\, P_{m-2} = P_m.$$

Also $B_{11} = 2\eta P_o P_{m-1} - P_o P_{m-2} = P_m$ (48b)

$$B_{mm} = 2\eta\, P_{m-1}\, P_o - P_{m-2}\, P_o = P_m. \qquad (48c)$$

Since N is symmetric we need show only the upper matrix terms ($j > i$) as zero:

($j > i$, $i \neq 1$)

$$B_{ij} = (-1)^{i+j}[-P_{i-2}P_{m-j} + 2\eta P_{i-1}P_{m-j} - P_i P_{m-j}]$$

(48d)

$$= (-1)^{i+j}P_{m-j}[-P_i + (2\eta P_{i-1} - P_{i-2})] = 0$$

($j > i$, $i = 1$)

$$B_{1j} = 2\eta(-1)^{1+j}\, P_o P_{m-j} + (-1)^{2+j}\, P_1\, P_{m-j}$$

(48e)

$$= (-1)^j\, P_{m-j}[-P_1 + (2\eta)] = 0$$

Thus, eqs. (48) verify (46).

Next, consider the existence of a value of $\eta \geq 2$ such that there are no extraneous inflection points. The second derivative at i, $y''_i$, can be expressed in two equivalent (cf. (5)) forms, (26) and (27). By adding these two forms we obtain the more symmetric expression:

$$2y''_i = \frac{1}{(\eta^2-1)} \frac{p^2 \sinh p}{\sinh p - p}[y'_{i-1} - y'_{i+1} + (\eta+1)(y_{i+1} - 2y_i + y_{i-1})] \quad (i = 1,\ldots,n) \quad (49)$$

or, using (22)

$$y''_i = \frac{1}{2(\eta^2-1)} \frac{p^2 \sinh p}{\sinh p - p}[y'_{i-1} - y'_{i+1} + (\eta+1)d_i] \quad (i = 1,\ldots,n) \quad (49')$$

Substituting (49') into (45) and omitting the positive constants, we must show that

$$[y'_{i-1} - y'_{i+1} + (\eta+1)d_i]d_i > 0 \quad (i = 0,\ldots,n+1) \quad (50)$$

for some $\eta \geq 2$.

Assume $d_i > 0$ and consider the $i^{th}$ inequality of (50) $(i \neq 0,n+1)$:

$$(\eta+1)d_i > y'_{i+1} - y'_{i-1}. \quad (51)$$

Using the solution for $y'_i$ (cf. (44)) where $N^{-1}$, (35), is expressed in terms of the elements of its adjoint (38), the inequality (51) becomes:

$$(\eta+1)d_i\, P_n \geq \sum_{j=1}^{n} (R_{i+1\,j} - R_{i-1\,j})C_j \quad (52)$$

From (36) we see that $P_i$ is an $i^{th}$ order polynomial in $\eta$; from (36) and (38'), $R_{ij}$ $(i,j = 1,\ldots,n)$ is a polynomial of order

n-1 or less; from (34), $C_1$ is linear in $\eta$. Further, it can be shown from (36) that

$$0 < P_o < P_1 < P_2 < \ldots < P_m, \quad \eta \geq 2. \tag{53}$$

Thus, we find that the R....S. of (52) is a polynomial in $\eta$ of order n while the L.H.S. of (52) is a polynomial in $\eta$ of order n+1 and is positive for $\eta \geq 2$. Thus, there will always be a value of $\eta \geq 2$ which satisfies the inequalities (52). Such a value will be the largest root of a polynomial of order n+1 and can be easily found numerically if not analytically. Other cases in the proof follow essentially this same argument.

Having determined the minimum value of $\eta$ which satisfies the curvature conditions, the unknown slopes can be determined from (44); (23) is then used to perform the interpolation in the desired regions.

An example of a "tensioned" spline curve is shown as a broken line in Figure 5. The solid line shows a normal spline curve through the 3 given points 0, 1.0, 2.0. The two intervals shown are actually part of a 7 interval problem; the vertical scale is exaggerated 5 times for clarity. Note that two extraneous inflection points are removed by the use of the spline in tension.

## 4.4 Some Practical Considerations

In the case where three consecutive points are in a straight line, i.e. $d_1 = 0$, the conditions (45) cannot be satisfied for finite $\eta$. An infinite value of $\eta$ will make the

interpolation curve simply a polygonal function - which is hardly satisfactory. Practical considerations suggest prescribing a straight line from $y_{i-1}$ to $y_{i+1}$ and considering two problems: a curve from $y_0$ to $y_{i-1}$ with $y'_{i-1}$ specified and a curve from $y_{i+1}$ to $y_{n+1}$ with $y'_{i+1}$ specified. Where $d_i$ is very close to zero, a circular segment prescribed through the given points at $y_{i-1}$, $y_i$, $y_{i+1}$ and the use of the above scheme would probably be preferable to a very large value of $\eta$. In general, specification of a discontinuity in curvature or slope at an intermediate given point, or an interval of specified curvature, can be handled by considering two spline curves, one on either side of the specified discontinuity or region.

If a predominance of problems deal with specified end curvatures, it is useful to derive the solution in terms of second derivatives at the given points instead of the slopes since it avoids the explicit use of the two additional equations (29). In such a change of variables, (4) rather than (5) are the conditions which are not identically satisfied, but the resulting matrix, N (cf. 32)), is unchanged.

## References

1. G. Birkhoff and H. L. Garabedian, "Smooth Surface Interpolation", J. Math. & Phys., 39 (1960), pp. 258-268.
2. I. J. Schoenberg, "Approximations of Equidistant Date by Analytic Functions", Quart. Math., 4 (1946), pp. 67-78.
3. I. J. Schoenberg and Anne Whitney, "On Polya Frequency Functions, III. The Positivity of Translation Determinants with an Application to the Interpolation Problem by Spline Curves", Trans. Amer. Math. Soc., 74 (1953), pp. 246-259.
4. I. J. Schoenberg,"Spline Functions, Convex Curves and Mechanical Quadrature", Bull. Amer. Math. Soc., 64 (1958), pp. 352-357.
5. R. S. Johnson, "On Monosplines of Least Deviation", Trans. Amer. Math. Soc., 96 (1960), pp. 458-477.
6. J. L. Walsh, J. H. Ahlberg and E. N. Nilson, "Best Approximation Properties of the Spline Fit", J. Math. & Mech., 11 (1962), pp. 225-234.
7. J. H. Ahlberg and E. N. Nilson, "Convergence Properties of the Spline Fit", J. Soc. Indust. Appl. Math., 11 (1963), pp. 95-104.
8. S. Timoshenko, "Strength of Materials, Part II", D. Van Nostrand Company, Inc., Princeton, N. J., Third Ed., 1956, pp. 41-46.
9. O. Taussky, "A Recurring Theorem on Determinants", Amer. Math. Monthly, 56 (1949), pp. 672-676.
10. D. E. Rutherford, "Some Continuant Determinants arising in Physics and Chemistry", Proc. Roy. Soc. Edinburgh, Series A, 62 (1947), pp. 229-236.
11. National Bureau of Standards, "Tables of Chebyshev Polynomials", Applied Mathematics Series 9, U. S. Gov. Print. Office, Washington, D. C., 1952.
12. T. Muir, "The Theory of Determinants in the Historical Order of Development, Vol. IV", Macmillan and Co., Ltd., London, 1923, p. 404.

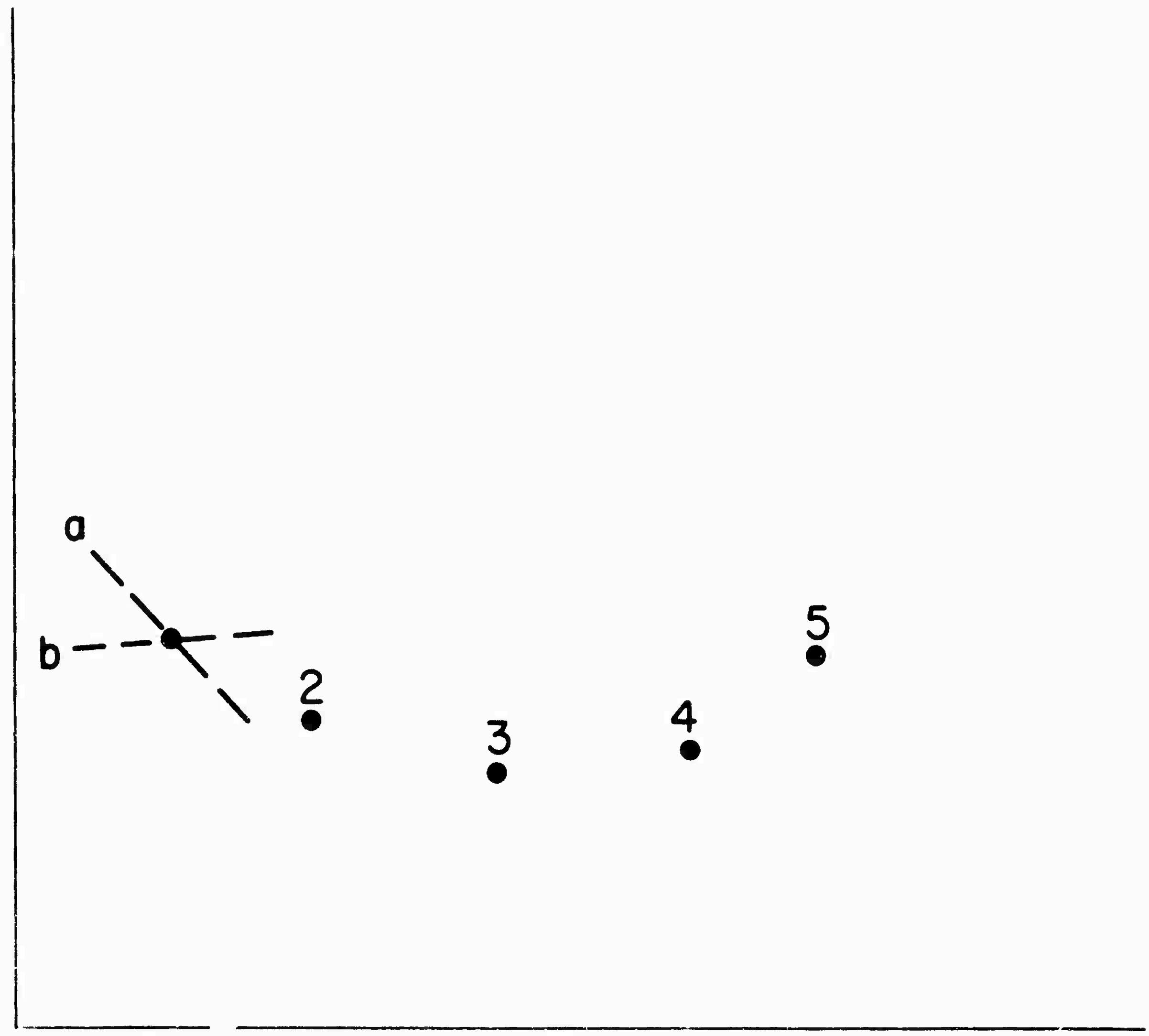



FIGURE 1

POINTS WITH POSITIVE SECOND DIFFERENCES

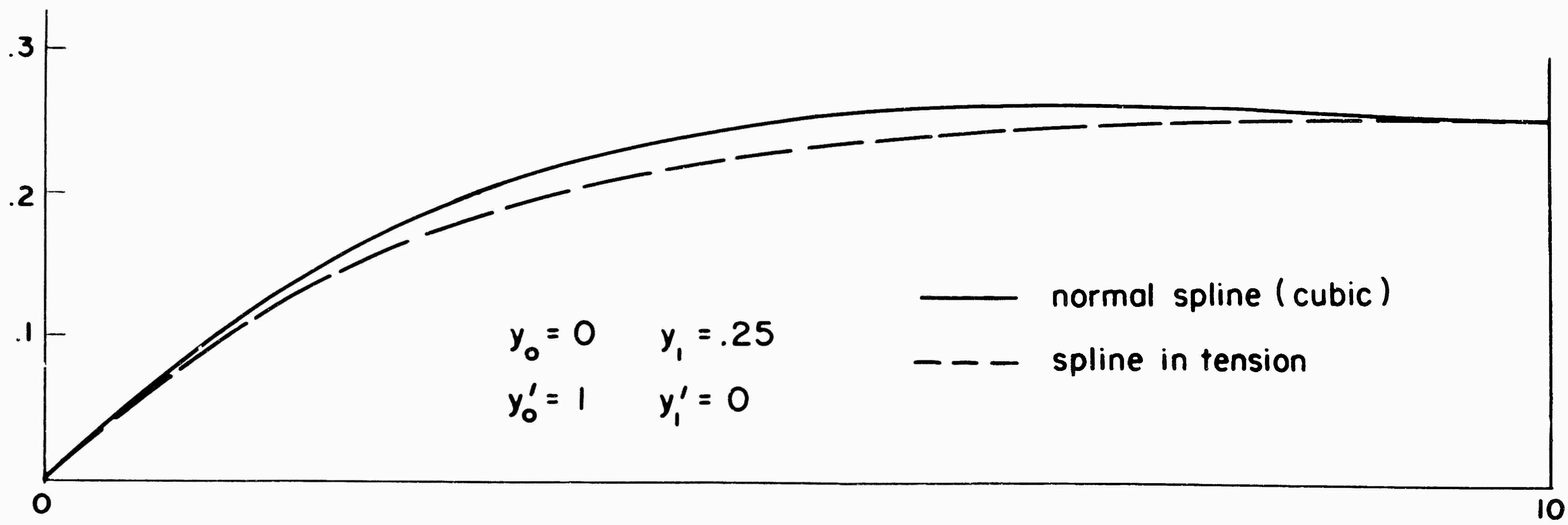
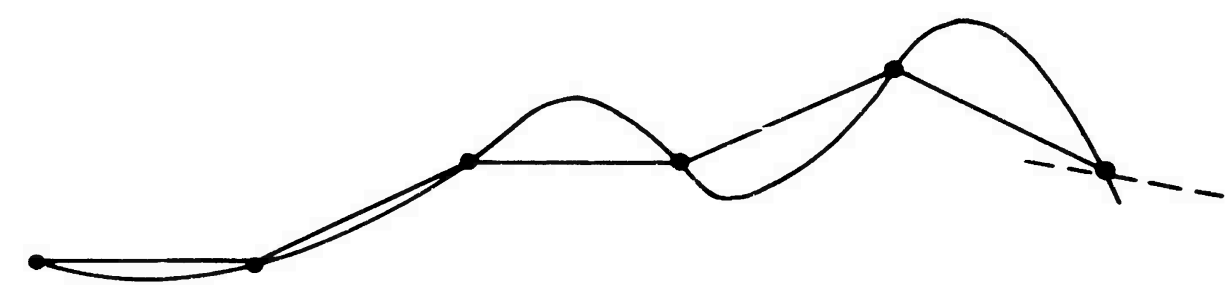

.3
.2
.1
0
10
$y_0 = 0$
$y_1 = .25$
$y_0' = 1$
$y_1' = 0$
normal spline (cubic)
spline in tension

FIGURE 2



| | | | | | | |
|---|---|---|---|---|---|---|
| $k =$ | $i-1$ | $i$ | $i+1$ | $i+2$ | $i+3$ | $i+4$ |
| sign of $y''_k =$ | | $+$ | $+$ | $-$ | $+$ | $-$ |
| sign of $d_k =$ | | $+$ | $-$ | $+$ | $-$ | $+$ |

FIGURE 3

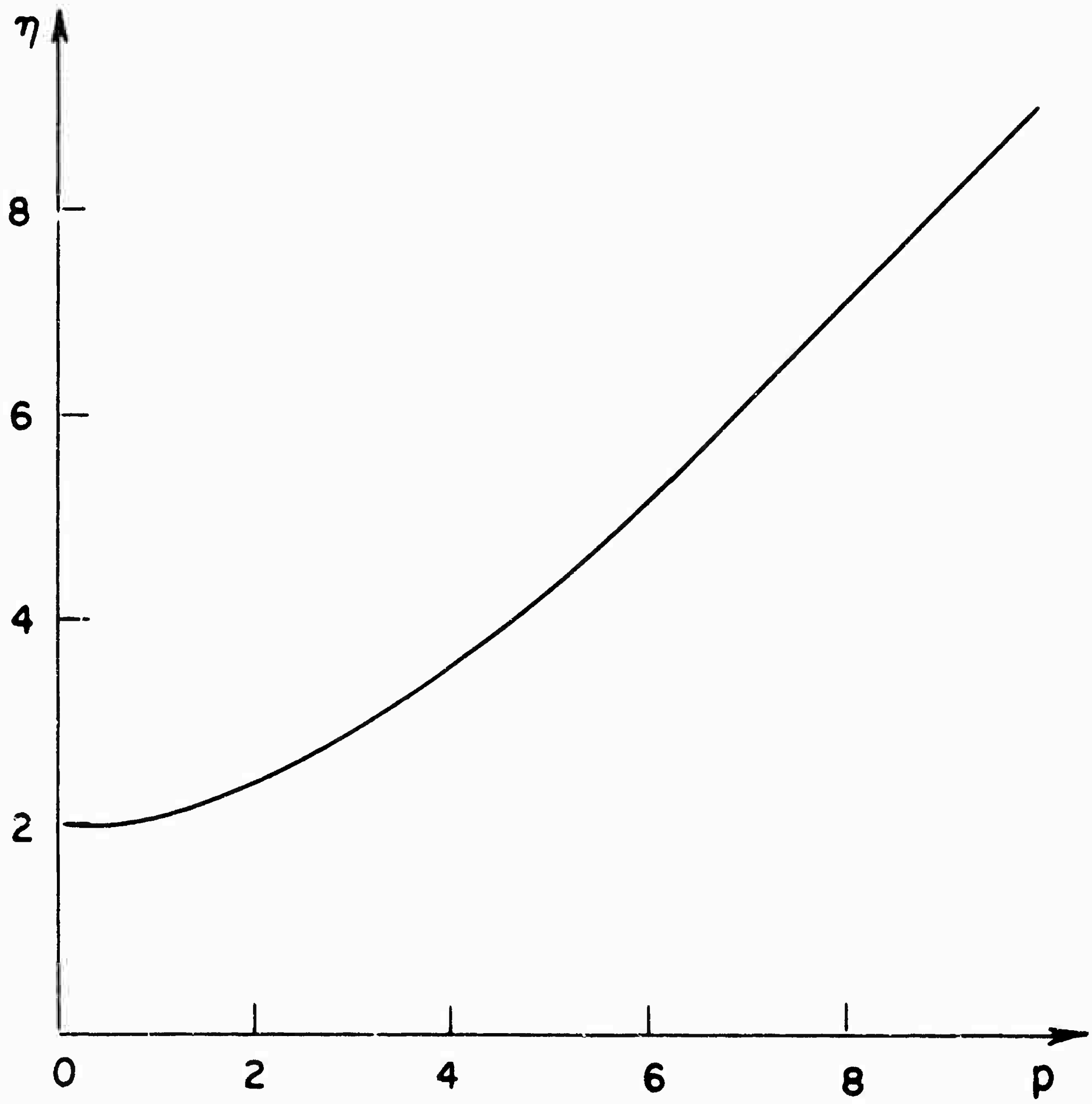



FIGURE 4

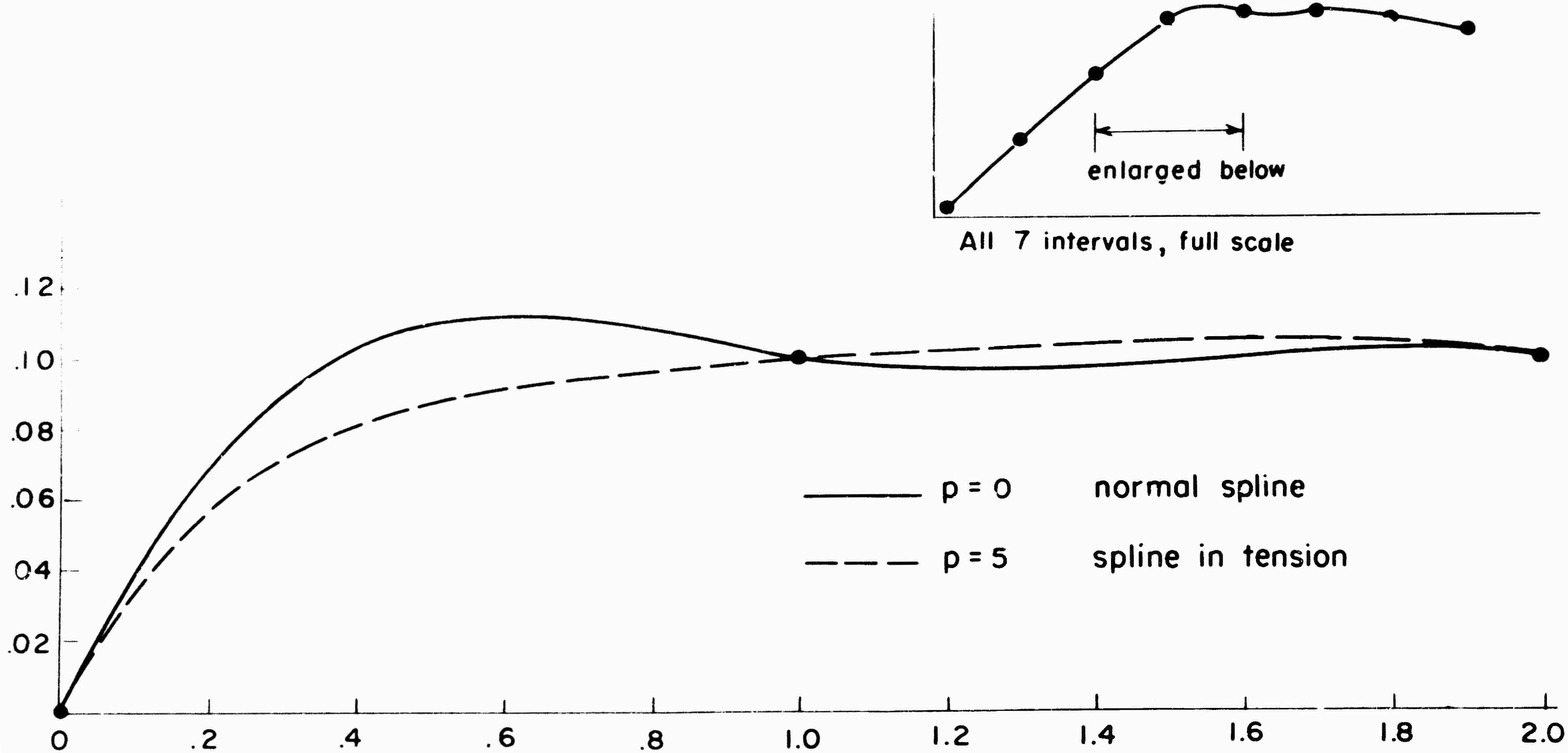

enlarged below
All 7 intervals, full scale
.12
.10
.08
.06
.04
.02
0
.2
.4
.6
.8
1.0
1.2
1.4
1.6
1.8
2.0
p = 0 normal spline
p = 5 spline in tension

FIGURE 5